DOCUMENT RESUME

ED 201 274 HE 013 753

AUTHOR Mizener, Deborah A.; Abrami, Philip C.
TITLE The Relationship Between Student/Teacher Attitude Similarity and Ratings of Instructional Effectiveness.
PUB DATE Apr 81
NOTE 33p.; Paper presented at the Annual Meeting of the American Educational Research Association (Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Attitude Measures; College Students; *Congruence (Psychology); Higher Education; *Interpersonal Attraction; Predictor Variables; *Student Attitudes; *Student Evaluation of Teacher Performance; *Student Reaction; Student Teacher Relationship; *Teacher Attitudes; Teacher Effectiveness

ABSTRACT

The relationship between student/teacher attitude similarity and teacher evaluations was studied in a group of 222 students who completed a 24-item attitude scale, once for themselves and once for their instructors, and a 28-item teacher rating form (TRF). A significant, moderate-sized correlation between assumed similarity and TRF scores for students pooled across classes was found. However, the size of the relationship dropped dramatically once the relationship was examined with all possible teacher effects removed. When controlling for teachers, students who perceived their own attitudes as similar to the instructor's assigned only slightly higher TRP scores than attitudinally dissimilar students. Analysis of component scores revealed that the set of 24 self-rating variables and the set of 24 perceived rating variables explained more TRF variance than the set of 24 individual similarity scores. Only one of the two sets of component scores contributed significantly to the prediction of TRF scores independent of all other effects. The relationship between perceived teacher attitudes and TRF scores was fairly substantial, and it is concluded that the perceived scores appear to be the best predictor of ratings. However, it is impossible to determine whether student differences in their perceptions of teacher attitudes affect their ratings of teacher effectiveness or vice versa. Theories of similarity and attraction and college classroom similarity research are reviewed, and a sample attitude scale is appended. (SW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED201274

The Relationship Between Student/Teacher Attitude Similarity And Ratings of Instructional Effectiveness

Deborah A. Mizener

McGill University, Montreal, Canada

Philip C. Abrami

Concordia University, Montreal, Canada

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. A. Mizener

P. C. Abrami

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

HE 013 753

Running head: Attitude Similarity

Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, California, April 1981.

Requests for reprints should be addressed to:

Deborah Mizener
Department of Educational Psychology
McGill University
3700 McTavish St.
Montreal, Quebec
Canada H3A 1Y2

Abstract

This study examined the relationship between student/teacher attitude similarity, the components of an assumed similarity score, and teacher rating form (TRF) scores. Students (N=222) completed a 24-item attitude scale, once for themselves and once for their instructors, and a 28-item TRF. The results indicated that TRF variance explained by total similarity was reduced considerably when teacher effects were removed. The component scores (self-ratings, perceived ratings) accounted for more TRF variability than assumed similarity. Of the component scores, only perceived ratings accounted for a significant amount of TRF variance independent of other effects.

The Relationship Between Student/Teacher Attitude Similarity And Ratings of Instructional Effectiveness

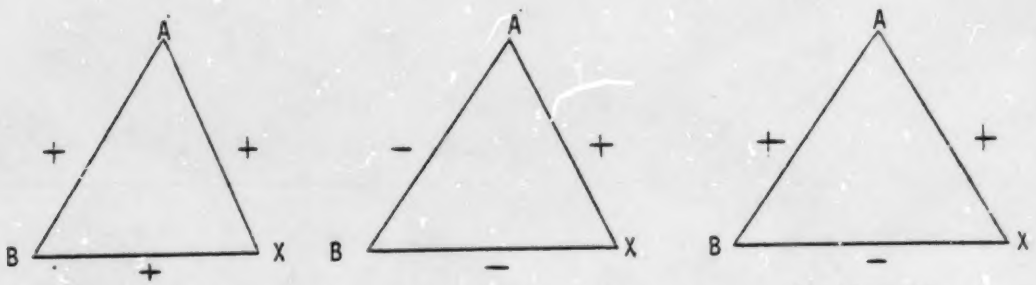
The implementation of teacher rating forms (TRFs) as instruments to measure teaching effectiveness has triggered much research to establish their validity. Whereas some research suggests that teacher evaluations do represent instructor impacts on student learning (Centra, 1977; Sullivan & Skanes, 1974), other findings suggest that ratings of instructors are correlated with student characteristics such as attitudes, values, and beliefs (Good & Good, 1973a, 1973b; Levenson & LeUnes, 1974; Hofman & Kremer, 1980). An interest in what ratings predict as well as how students generate their evaluative responses is an interest in the convergent-discriminant validity (Campbell & Fiske, 1959) of teacher ratings. Thus, for teacher ratings to possess discriminant validity, they should be uninfluenced by irrelevant factors, such as the characteristics of the student raters and perhaps, the similarity of students and teachers.

Interpersonal similarity refers to the state of being alike; it is synonymous with interpersonal congruence, match or fit as we use those terms here. For centuries, it has been a popular notion that people tend to like individuals who they perceive as similar to themselves on various attributes. Writers and philosophers such as Aristotle, Spinoza, and Swift (Byrne, 1971) have commented on the unpleasantness of interactions that involve the exchange of dissimilar viewpoints. The expression of similar views, however, tends to be associated with positive interpersonal responses. In consort with this view, critics of teacher rating forms have speculated that the influence of student/teacher attitude similarity may be large enough to generate substantial "bias" in teacher ratings.

Theories of Similarity and Attraction

Attitude similarity denotes the actual or perceived likeness of attitudes between individuals. Two major theoretical viewpoints have been proposed to explain how

attitude similarity may lead to interpersonal attraction and thus affect the teacher ratings made by students. Of the cognitive-consistency theories, balance theory (Festinger, 1957; Heider, 1958; Newcomb, 1959) provides a sufficient explanation. Basically, people prefer relationships that fit well together - are balanced - than relationships that are discordant. Since balance is a psychologically preferred state, individuals will strive for consistency in relationships. Thus, when Person A and Person B agree concerning X (which could be any person, object or attitude statement), balance is possible (for A) only if A likes B. On the other hand, if A and B disagree concerning X, balance occurs when A dislikes B. For example, a state of balance would exist if Person A was positively attracted to Person B, held a positive attitude toward X, and perceived that B also held a positive view of X. This would produce a psychologically comfortable state; however, a discrepancy between the attitude held by Person A and Person B would lead to a state of imbalance. This is only the case if A likes B. A discrepancy between the attitude held by Person A and Person B toward X would lead to a balanced state if A dislikes B. The following diagram indicates balanced and imbalanced states:



According to Newcomb (1961), life in a balanced state is easier due to the facilitation of communications and interactions. His classic study (Newcomb, 1961) demonstrated the relationship between attitude similarity and attraction. He investigated how men residing in university dormitories developed friendships. He

periodically measured their attitudes on a variety of topics and assessed their attraction for one another. The results indicated that individuals tended to assume that those they liked shared their attitudes on important and relevant matters. It was also found that as the individuals became familiar with the attitudes of others, actual attitude congruence did result in attraction.

An alternative explanation of attitude similarity and attraction has been proposed by reinforcement theorists (Byrne, 1969; Byrne & Clore, 1970; Clore & Byrne, 1974). Unlike the balance theorists who focus on cognitions, these theorists tend to focus on stimuli and responses as the basic units of interpersonal interactions. According to Byrne (1971), the balance and reinforcement models do not constitute alternative and mutually exclusive explanatory systems. The attitude-attraction function is interpretable in either cognitive or reinforcement terms. But these two interpretational schemas utilize different languages and lead to different types of empirical research.

Golightly and Byrne (1964) devised a learning task in which similar and dissimilar attitude statements were used as reinforcers. When subjects responded correctly to the discrimination task, they received a card with an attitude statement which agreed with their position. When they responded incorrectly, they received a disagreeing statement. The probability of the occurrence of a response increased when the response was followed by a similar attitude statement and decreased when followed by a dissimilar attitude statement, indicating that similar and dissimilar attitude statements can act as positive and negative reinforcers, respectively.

A person who displays similar or dissimilar opinions to an individual may be viewed as supporting or opposing the subject's own views. Byrne and Clore (1967) suggest that such "consensual validation" is often a rewarding experience and that disagreements are often experienced as punishing. In the absence of other information, consensual validation helps determine whether one is logical or correct in interpreting environmental events. When people provide validation by indicating that their percepts and concepts are congruent with others, rewarding interactions occur and, hence,

positive relationships develop. Dissimilarity of ideas would lead to punishing interactions and as a result, negative relationships might develop.

Byrne (1961) tested the proposition that attitude similarity would lead to attraction with the attraction-to-a-stranger paradigm (also used in subsequent studies). A 26-item broadband attitude questionnaire was administered to introductory psychology students. Two weeks later, under the pretext of an interpersonal prediction study, students were asked to form judgments about a stranger after reading the stranger's responses to the same attitude scale. In reality, the experimenter had completed the questionnaires and distributed them so that some subjects received scales in which the stranger's responses were exactly similar to those of the subjects (similar stranger), while others received scales with reponses exactly opposite to their own (dissimilar stranger). Byrne's hypothesis was confirmed: the subjects indicated greater liking for similar strangers than dissimilar strangers. The similar strangers also received higher ratings on intelligence, knowledge of current events, morality, and adjustment.

Byrne and Nelson (1965) tested the proposition that attraction toward a stranger is a positive function of the proportion of positive reinforcements received from that stranger. They found a linear relationship between proportion of similar attitudes and attraction and proposed a tentative law of attraction which maintains that "attraction toward X is a positive linear function of the proportion of positive reinforcements received from X" (Byrne & Nelson, 1965, p. 662). Based on this and other research, Byrne and Clore (1970) detail further this associational theory which is the basis of evaluative responses. This relationship has been verified with children (Byrne & Griffith, 1966) as well as Job Corps trainees and hospitalized schizophrenics (Byrne, 1969). The nonlaboratory generalizability of the attraction relationship was examined in a computer dating field study (Byrne, Ervin, & Lamberth, 1970) where it was found that attraction was significantly related to attitude similarity as well as physical attractiveness.

College Classroom Similarity Research

Researchers (Good & Good, 1973a, 1973b; Levenson & LeUnes, 1974; Hofman & Kremer, 1980) have investigated the relationship between student/teacher attitude similarity and student ratings of college instructors. They have generally found what appears to be a positive relationship between attitude similarity and student ratings of instruction. For example, if a student and teacher from a political science course prefer the same political party, then the student's evaluation of that teacher should be higher than that of a student who is dissimilar to the teacher in political party preference.

Most of these attitude similarity studies deal with actual and/or assumed attitude similarity. Actual similarity typically refers to the difference between a student's self-rating and a teacher's self-rating. Assumed or perceived attitude similarity is the difference between a student's self-rating and a student's rating of the instructor's perceived attitudes.

Good and Good (1973a) hypothesized that subjects with similar attitudes to an instructor would assign higher student ratings than students with dissimilar attitudes. To test their hypothesis, Good and Good had undergraduate students complete an abbreviated form (14 items) of Byrne's (1971) Survey of Attitudes. During a subsequent class session, each student received an attitude survey representing the attitudes and opinions of a hypothetical college instructor who showed either 14% agreement (2 out of 14 items) or 86% agreement (12 out of 14 items) with the student. Overall, the similar group assigned higher ratings than the dissimilar group on items such as competence, open-mindedness, stimulation and likeability.

In a later study, Good and Good (1973b) investigated the relationship of assumed and actual attitude similarity with course evaluations. Fourteen instructors and their 409 students filled out a 24-item version of Byrne's (1971) Survey of Attitudes. Students completed the attitude questionnaire twice; once rating themselves and once rating the teacher's perceived attitudes. Assumed similarity showed moderately positive

correlations with several instructors and course evaluation variables. However, there appeared to be little evidence of any relationship between actual similarity and student evaluations of instruction.

Levenson and LeUnes (1974) had psychology students in a single class complete a TRF and a 14-item Byrne (1971) Survey of Attitudes. The instructor also completed the attitude scale. They found that actual attitude similarity was related to how students felt personally about an instructor, however, similarity was not related to ratings of an instructor's competence.

Hofman and Kremer (1980) conducted two studies to explore the relationship between course evaluations and the attitudes of students and instructors toward higher education. (Unlike Byrne's scale, their attitude scale focuses uniquely on college teachers and instruction and not on an array of topics.) The actual similarity index was the reciprocal of the D statistic (Osgood, Suci, & Tanenbaum, 1957, p. 91) summed over the squared differences between student and instructor scores on the 24 attitude items. When rating scores were regressed on student attitudes, teacher attitudes, and similarity indices, 17.7% of the rating variance was accounted for (6.2% when similarity was entered first into the analysis and 2.6% when similarity was entered last). The second study was modified slightly: students indicated their perceptions of their instructor's attitudes while the instructors indicated their perceptions of the average student's attitudes. Four attitude sets (students' own attitudes (SO), instructor's own attitudes (IO), students' attitudes as perceived by instructors (SP), and instructor's attitudes as perceived by students (IP)) were created along with four similarity indices (SO and IO, SO and SP, IO and IP, SP and IP). A stepwise multiple regression analysis acounted for 50% of the variance once the rating scores were regressed on the item sets and the set of congruences. The congruences alone accounted for 13% of the rating variance. Hofman and Kremer (1980) concluded that course evaluations are related to student and instructor attitudes toward higher education as well as the similarity between these attitudes. Ratings of instruction will increase when students

perceive their instructors as possessing certain attitudes toward higher edcuation and teaching. Furthermore, they suggest that teacher ratings are limited in their ability to reflect differences in teacher effectiveness and their validity must be questioned.

Collectively, these data suggest a relationship between student/teacher attitude similarity and ratings of instructional effectiveness. But we are not convinced that this evidence is sufficient to claim that attitude similarity is either a substantial or frequent source of bias in student ratings. First, there are methodological problems and inconsistencies across the studies which cloud the issue including: failure to separate teacher-produced rating variability from student-produced variability, variable operational definitions of similarity, and the unit of analysis problem. Second, Grush, Costin and Clore's (1975) extension of Byrne's law of attraction suggests conditions under which student/teacher similarity will not lead to higher evaluations.

An important consideration in TRF validity research is the potential confounding of student- and teacher-produced rating variability. In order to explore student effects and exclude variability produced by teachers, the student within a class can serve as the unit of analysis as in Levenson and LeUnes (1974) who studied the students of only one hypothetical instructor. To study teacher effects, class means are usually appropriate units of analysis particularly when class-to-class differences in students are controlled through randomization. The problem occurs when students from two or more classes are pooled thus incorporating variability due to students with variability attributable to between class differences in instructors. Such is the case with Good and Good (1973b), and Hofman and Kremer (1980). It seems important to us that before one can make a meaningful statement about "bias" in ratings, one has to control, at a minimum, for the substantial variability in ratings which is legitimately attributable to teacher effects.

The second problem which arises is the diverse array of operational definitions of attitude similarity. Good and Good (1973a) studied the effects of actual similarity scores as a manipulated variable, while Levenson and LeUnes (1974) studied actual

similarity as a naturally occurring variable. The studies also differed in defining a similarity criterion (agreement on 12 out of 14 statements as opposed to 10 out of 14, respectively). Furthermore, Levenson and LeUnes (1974) discarded more than two-thirds of their students from the final analysis. Neither of these criteria was used by Good and Good (1973b) or Hofman and Kremer (1980) who employed the constant-discrepancy technique and D statistic and measured various similarity indices in real classroom situations. The result is a vaudeville of similarity scores which are not highly comparable, yet which might speak to the generalizability and robustness of the phenomenon.

We disagree. The differences in operational definitions of similarity in consort with differences in analysis procedures do not add support to the phenomenon but detract from it. In particular, these data do not lend equal weight to the interpretation that teacher ratings lack discriminant validity. One must be cognizant of the sources of variability each similarity score can represent. Consider, for example, the difference between actual and assumed similarity. An actual similarity score is the difference between a student and instructor characteristic; an assumed similarity score reflects two student characteristics and some impact of instructor characteristics on student perceptions. When computed within class, as in Levenson and LeUnes (1974), an actual similarity score reflects only student-produced variability since teacher effects are constant. When computed across classmeans, an actual similarity score reflects only teacher effects since student effects are controlled by randomization. On the other hand, when computed within a class an assumed similarity score reflects two sources of student variability - their own attitudes and their perceptions of the instructor - and not teacher effects. However, when computed across classmeans, an assumed similarity score reflects only teacher-produced differences in student perceptions. Obviously, an analysis of either actual or assumed similarity based on pooled data will combine variability due to students and teachers. Therefore, we view as contentious the suggestion that all attitude similarity/ratings relationships provide evidence that ratings

are biased. We remain unconvinced unless teacher effects on student attitudes and student ratings are controlled. The bulk of prior research fails to provide these controls.

Another issue concerns similarity and Byrne's law of attraction. Grush et al. (1975) extended Byrne's law of attraction to include situations where positive dissimilarity between student and instructors on relevant personality traits was positively related to ratings of instructional effectiveness. That is, instructors who were perceived as having more of a desirable, situationally relevant trait than students were rated higher than instructors who were seen as having less of that characteristic but yet were similar to students. Grush et al. (1975) suggest that students engage in social comparison when making their teacher evaluations. Students are attracted to instructors who possess more of certain traits and values than the students possess because of the belief that learning is facilitated by instructors who know more or have more desirable traits than students. Thus, it is not the similarity of students and teachers which is reinforcing but certain instructor characteristics as perceived by students which are reinforcing. Such an analysis suggests that the relationship between student/teacher attitude similarity - in particular assumed similiarity - needs to be examined more closely. Rather than being a function of assumed similarity, variance in ratings may be explained by student perceptions of instructor attitudes.

Hofman and Kremer's (1980) findings also suggest the importance of examining component scores: student perceptions of teacher attitudes accounted for a substantial amount of rating variance - more of the variance than any other source. Unfortunately, Hofman and Kremer (1980) administered their highly relevant attitude scale at midsemester, giving the instructors an ample opportunity to influence student attitudes. It is not surprising, then, that students' own attitudes and their perceptions of instructor attitudes are substantially related to ratings. But this is far from sufficient evidence to claim "bias" in student ratings.

The main purpose of the present research was to explore more carefully the

relationship between attitude similarity and teacher evaluations. More specifically we had three concerns. The first concern was to determine whether the relationship between assumed attitude similarity and TRF scores was meaningfully reduced once teacher effects were considered. Second, we wondered whether the components of an assumed similarity score accounted for substantially more TRF variability than similarity scores. The final purpose was to examine which of the component scores was the best predictor of TRF scores. Our concern was not with theory testing but whether the theory has a practical use in a field setting.

Method

Attitude scale construction

Initially, we considered adopting Byrne's Survey of Attitudes (reprinted in Byrne, 1971) for use in our research. The scale was created in 1960 by asking undergraduates to list issues which they discussed with their friends; the 26 most frequently mentioned topics were converted into a scale consisting of 56 six-choice items. The test-retest reliability of these attitude items ranged from .55 to .95. Upon reviewing the scale, however, a large number of items appeared obsolete and inappropriate. Consequently, we decided to create a new scale following Byrne's procedures and utilizing his items wherever possible.

Phase I. In Winter, 1979, a questionnaire was distributed to 39 education students enrolled in a compulsory graduate statistics course at McGill University. Students were told: "We are constructing an attitude scale to be used for research purposes. We are interested in including items on the scale which stem from important current topics. It would be appreciated if you would list a series of issues which you and your acquaintances discuss." Students could list an indefinite number of topics, however, they were constrained by a 15-minute time limit. Student responses were roughly categorized into five areas: interpersonal relationships, recreation, educational concerns, politics and its influences, and miscellaneous. The number of responses per category was 65, 59, 50, 97, 33, respectively. A proportional number of items were created per

category or retained from Byrne. Of Byrne's original 56 items, 12 were revised, 22 slightly changed, and five omitted. The items omitted seemed either obsolete or inapplicable. For example, "Necking and Petting and "Freshmen Having Cars on Campus" were discarded items. The slight changes were to allow for consistency amongst the items i.e., like and dislike instead of enjoy and dislike, while other items were revised so they would be equally meaningful to Canadians as Americans. Finally, 17 new items were added creating a scale of 69 items.

Phase II. A second series of studies was conducted to test for item ambiguity and topic importance. During the 1980 winter session, 40 undergraduate art students at McGill University were asked to rate "...the extent to which the items and their alternatives address issues of importance to you." They rated each item on a 5-point scale ranging from extremely important (1) to not at all important (5). Another 40 undergraduate education students were asked to rate on a five-point scale "...the extent to which the items and their alternatives are clear, understandable, and unambiquous." All students were informed of the study's purpose: to construct a questionnaire to measure student attitudes toward a variety of topics. Item means and standard deviations were tabulated for both topic importance and ambiguity. In order to condense the scale, items which were ambiguous ($x<2.5$) and unimportant ($x>3.5$) were discarded. Five new items which otherwise met these criteria were discarded because they were specific to Quebec (Sovereignty Association) or seemed repetitive of Byrne items. For example, "Women's Liberation Movement" would probably be highly related to "Careers for Women." The final attitude scale (see Appendix) consisted of 24 items; 9 new ones, 4 original Byrne items, and 11 revised Byrne items (6 of which had minor changes). The mean importance rating of the 24 final scale items was 2.090; the mean ambiguity rating was 3.9708. (See Table 1.)

Phase III. To assess the stability of attitude scale responses, test-retest reliabilities were computed based on data collected in the Fall, 1980. Thirty-seven students enrolled in education courses at McGill completed the revised Survey of

Attitudes on two occasions, three weeks apart. These students had not participated in other phases on the project. On each occasion the students were given 15 minutes to complete the form. At the second session, the instructions mentioned "...people's attitudes sometimes change, therefore, please read each item carefully and choose the one alternative which presently describes your personal attitude on that topic."

The 24 reliability coefficients are reported in Table 1. They ranged from +.06 to +.88 averaging .60. The average for the 15 Byrne items was .56 while the average for the newly created items was .67. These reliabilities are lower than those reported by Byrne (1971) and may be so for two reasons. First, midterm exams occurred between testing sessions which may have lowered the coefficients of university-related items in particular. Second, we tested only a small, rather homogeneous group of students. Had we used a larger, more diverse student sample, the reliabilities might have increased substantially. Consequently, we were especially pleased that relative to the Byrne items the newly created items had high reliabilities, although somewhat lower than we originally expected.

Similarity study

Subjects. The subjects for the main study were 222 undergraduate McGill University students enrolled in 14 classes taught in the Faculty of Education during the 1979-1980 academic year. Students volunteered to participate in the study and were not fully informed of its purpose until the conclusion of data collection when they were debriefed.

Instruments. The 24-item attitude scale was used to assess student self-ratings and student perceptions of instructor attitudes. The TRF consisted of 28 items, each rated on a 5-point scale. Three questions represented global items of teaching effectivenes. The remaining items were adapted from Hildebrand and Wilson (1970) and asked about specific aspects of teaching.

Procedure. Questionnaires were administered at midsemester over three class meetings. First, students rated their instructor on the TRF. Second, students

completed the 24-item attitude scale, rating their own attitudes. Finally, each student completed the attitude survey again but giving their impressions of the instructor teaching the course.

Data analysis. Student own attitude responses to the scale were factor analyzed using orthogonal, varimax rotation via subprogram FACTOR in SPSS (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975). A principal components analysis revealed 10 factors which accounted for 24.0%, 15.2%, 11.7%, 9.2%, 8.4%, 7.8%, 7.4%, 7.3%, 5.2%, and 3.9%, respectively, of the scale variance. Since a large number of factors emerged, this suggested to us that the scale does indeed address a diversity of topics. Consequently, we chose to avoid further analysis of factor scores; we concentrated, instead, on individual item analysis, where appropriate.

For each of the 24 attitude items, an assumed similarity score was the difference between student self-ratings and student perceptions of teacher attitudes. A total similarity score was computed by adding the absolute differences of the assumed similarity scores on all the attitude items.

Two measures of each student's general impression of his/her instructor's effectiveness was computed: the average response to the first three global TRF items and the average response to the remaining 25 specific TRF items. (Only the results for the latter measure are reported here since the findings were substantially the same.)

Multiple regression analyses were performed using subprogram REGRESSION from the SPSS package of computer programs (Nie et al., 1975). The first analysis was the regression of total similarity on TRF scores. The second analysis, designed to determine the relationship between similarity and TRF scores independent of teacher effects, was a stepwise regression where teacher effects were entered first followed by total similarity. In this and subsequent analyses, dummy variable coding was used to create 13 vectors representing the different classes. The third and fourth analyses were designed to assess the predictive ability of the component parts (self-ratings and perceived ratings) of a similarity score. Stepwise regressions were performed where

teacher effects were entered on the first step and component scores alternately entered on the second or third step. Unlike a total similiarity score, it did not seem conceptually reasonable to compute a total score for either the own attitude or perceived attitude component scores. For these later analyses, we entered each component as sets of 24 individual scores corresponding to the number of attitude items.

Results

The means and standard deviations for student own attitudes, perceived teacher attitudes and the individual teaching rating variables are presented in Table 2. The simple correlation between total similarity and mean TRF scores was significant and moderately sized (r=.28, $p<.01$). But the variance in TRF scores accounted for by total similarity was greatly reduced when teacher effects on ratings were removed. When the teacher vectors were entered first in a stepwise regression, the variance in TRF scores due to total similarity was still significant but rather small (ΔR^2 =.03, $F(1,207)$=9.42, $p<.01$)). This represented a decrease of approximately two-thirds in TRF variance explained by total similarity. (See Table 3.)

The set of 24 individual similarity scores explained a significant amount of variance in TRF scores independent of teacher effects (ΔR^2=.13, $F(24,184)$=1.85, $p<.05$). However, the two sets of component scores combined explained almost twice the variance in TRF scores independent of teacher effects than the set of individual similarity scores (ΔR^2=.25, $F(48,160)$=1.90, $p<.01$). (See Table 4.)

Both sets of component scores did not add significantly to the prediction of the TRF scores. When the set of 24 self-rating variables was entered on the last step of a stepwise regression (following the teacher vectors and the set of perceived ratings), the TRF variance explained was not significant ($F(24,160)$=1.36, $p>.05$). However, when the set of 24 perceived rating variables was entered on the last step of a stepwise regression, the TRF variance explained was significant and large (ΔR^2 =.15, $F(24,160)$=2.30, $p<.01$). (See Table 5.)

Discussion

Consistent with prior research, we found a significant, moderate-sized correlation between assumed similarity and TRF scores for students pooled across classes. However, the size of the relationship dropped dramatically once the relationship was examined with all possible teacher effects removed. When controlling for teachers, students who perceived their own attitudes as similar to the instructor's assigned only slightly higher TRF scores than attitudinally dissimilar students. Thus, prior claims of bias in TRF scores due to the effects of student/teacher attitude similarity may have substantially overstated the case.

The analysis of the component scores revealed that the set of 24 self-rating variables and the set of 24 perceived rating variables explained more TRF variance than the set of 24 individual similarity scores. Therefore, in terms of predictive ability, it might be wiser in future to ignore similarity scores and concentrate on the components of a similarity score.

Furthermore, only one of the two sets of component scores contributed significantly to the prediction of TRF scores independent of all other effects. Even with the conservative variance estimate of the present analysis, the relationship between perceived teacher attitudes and TRF scores was fairly substantial. Consistent with prior findings with attitudes (Hofman & Kremer, 1980), and personality traits (Grush et al. 1975), the perceived scores appear to be the best predictor of ratings. As a single source of variability, this is inconsistent with the claim that attitude similarity affects ratings but consistent with Bryne's law of attraction as interpreted by Grush et al. (1975). But whether TRF variance explained by perceived teacher attitudes is a source of bias in TRF scores, is equivocal. In this and prior studies, it is impossible to determine whether student differences in their perceptions of teacher attitudes affect their ratings of teacher effectiveness or vice versa. It is conceivable to imagine students who form their judgments of an instructor's effectiveness and then "round out the picture" by forming a balanced view of the instructor's attitudes and values.

Finally, our results are limited by several factors suggesting the need for follow-up studies. First, the instructors volunteered to participate; we believe this sample may overly represent instructors who receive average or above average ratings. Second, the number of instructors and students sampled was small. Lastly, these data were collected primarily in one faculty of a large, prestigious and research-oriented university and may not generalize to other settings.

References

Byrne, D. Response to attitude similarity-dissimilarity as a function of affiliation need. Journal of Personality, 1962, 30, 164-177.

Byrne, D. Attitudes and attraction. In L. Berkowitz (Ed.), Advances in experimental social psychology. New York: Academic Press, 1969.

Byrne, D. Interpersonal attraction and attitude similarity. Journal of Abnormal and Social Psychology, 1961, 62, 713-715.

Byrne, D. The attraction paradigm. New York: Academic Press, 1971.

Byrne, D., & Clore, G. L. Effectance arousal and attraction. Journal of Personality and Social Psychology, 1967, 6(4, Whole No. 638).

Byrne, D., & Clore, G.L. A reinforcement model of evaluative responses. Personality: An International Journal, 1970, 1, 103-128.

Byrne, D., Ervin, C. R., & Lamberth, J. Continuity between the experimental study of attraction and real-life computer dating. Journal of Personality and Social Psychology, 1970, 16, 157-165.

Byrne, D., & Griffith, W. A developmental investigation of the law of attraction. Journal of Personality and Social Psychology, 1966, 4, 699-702.

Byrne, D., & Nelson, D. Attraction as a linear function of proportion of positive reinforcements. Journal of Personality and Social Psychology, 1965, 1, 659-663.

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 1959, 56, 81-105.

Centra, J. A. Student ratings of instruction and their relationship to student learning. American Educational Research Journal, 1977, 14, 19-24.

Clore, G. L., & Byrne, D. A reinforcement-affect model of attraction. In Ted L. Huston (Ed.), Foundations of interpersonal attraction. New York: Academic Press, 1974, 143-165.

Centra, J. A. Student ratings of instruction and their relationship to student learning. American Educational Research Journal, 1977, 14, 19-24.

Clore, G. L., & Byrne, D. A reinforcement-affect model of attraction. In Ted L. Huston (Ed.), Foundations of interpersonal attraction. New York: Academic Press, 1974, 143-165.

Festinger, L. A theory of cognitive dissonance. Stanford, California: Stanford University Press, 1957.

Golightly, C., & Byrne, D. Attitude statements as positive and negative reinforcements. Science, 1964, 146, 789-799.

Good, K. C., & Good, L. R. Attitude similarity and attraction to an instructor. Psychological Reports, 1973, 91, 285-290.

Grush, J. E., Clore, G. L., & Costin, F. Dissimilarity and attraction: When difference makes a difference. Journal of Personality and Social Psychology, 1975, 32, 783-789.

Heider, F. The psychology of interpersonal relations. New York: Wiley and Sons Publishing Co., 1958.

Hildebrand, M. & Wilson, R.C. Effective university teaching and its evaluation. Washington, D.C.: U.S. Office of Education, 1970.

Hofman, J. E., & Kremer, L. Attitudes toward higher education and course evaluation. Journal of Educational Psychology, 1980, 72(5), 610-617.

Levenson, H., & LeUnes, A. Students' evaluation of an instructor: Effects of similarity of attitudes. Psychological Reports, 1974, 34, 1074.

Newcomb, T. M. Individual systems of orientation. In S. Koch (Ed.), Psychology: A study of a science (Vol. 3). New York: McGraw-Hill Publishing Co., 1959.

Newcomb, T. M. The acquaintance process. New York: Holt, Rinehart and Winston, 1961.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. (Eds.). Statistical package for the social sciences (2nd ed.). New York: McGraw-Hill, 1975.

Osgood, C., Suci, G., & Tannenbaum, P. The measurement of meaning. Urbana-Champaign: University of Illinois Press, 1957.

Sullivan, A. M., & Skanes, G. R. Validity of student evaluation of teaching and the characteristics of successful instructors. Journal of Educational Psychology, 1974, 66, 584-590.

Table I

Scale Construction: Mean Attitude Scores for Topic Importance and Ambiguity, And Reliability Coefficients of the Final Scale Items[a]

	Item Stem	Topic Importance		Ambiguity		Reliability
		Mean	S.D.	Mean	S.D.	
1.	Professors and Student Needs*	2.025	1.097	4.000	1.132	.06
2.	Birth Control*	1.850	1.099	3.750	1.463	.80
3.	Sports**	2.400	1 215	4.000	1.261	.81
4.	Money**	2.175	0.984	3.825	1.375	.62
5.	Grades*	2.000	1.062	4.225	1.097	.24
6.	Political Orientation**	2.225	0.832	3.900	1.257	.57
7.	Welfare System**	2.150	0.893	3.825	1.412	.67
8.	Creative Work**	2.125	0.911	3.925	1.289	.86
9.	Socialized Medicine**	2.150	1.075	4.000	1.109	.79
10.	University Education**	1.725	0.784	4.000	1.155	.59
11	Exercise**	2.150	0.945	4.350	1.001	.62
12.	Discipline of Children**	1.975	0.733	3.925	1.366	.01
13.	Women in Today's Society**	2.000	0.877	3.725	1.132	.55
14.	Family Finances**	1.975	0.733	3.750	1.391	.40
15.	Careers for Women*	1.900	0.841	4.450	0.876	.81
16.	Sex Relations***	2.500	1.013	3.850	1.252	.74
17.	Institution of Marriage***	2.025	0.862	3.925	1.248	.73
18.	Travelling***	2.350	1.075	4.375	0.979	.88
19.	Revising University Level Education***	2.425	0.874	3.725	1.176	.56
20.	University***	2.175	0.931	3.900	1.482	.17

Table 1 continued

	Item Stem	Topic Importance		Ambiguity		Reliability
		Mean	S.D.	Mean	S.D.	
21.	Third World War***	2.175	1.318	4.075	1.248	.73
22.	Inflation***	2.150	1.145	4.425	1.035	.76
23.	World's Future***	1.500	0.716	3.575	1.430	.65
24.	Career-Mindedness***	2.029	1.014	3.800	1.181	.82
		N=40		N=40		N=37

Midterm examinations occurred between testing sessions. This may have lowered the reliability of some of the items, particularly university-related ones (i.e., numbers 1, 5, and 20).

*Original Byrne item
**Revised Byrne item
***New item

Table 2

Means and Standard Deviations of Teacher Ratings and Attitude Items

(N=222)

	Attitude Items	Own Attitudes		Perceived Attitudes	
		Mean	S.D.	Mean	S.D.
1.	Professors and student needs	4.06	1.01	4.58	1.11
2.	Birth control	2.20	0.08	2.72	1.39
3.	Sports	1.77	0.89	2.29	0.93
4.	Money	2.60	1.24	2.63	1.18
5.	Grades	3.17	1.38	2.93	1.54
6.	Political orientations	2.32	0.96	2.50	1.37
7.	Welfare system	3.65	1.38	3.57	1.45
8.	Creative work	1.74	0.81	1.84	1.11
9.	Socialized medicine	4.92	1.16	4.69	1.15
10.	University education	3.38	1.38	2.20	1.21
11.	Exercise	5.14	0.91	4.68	1.28
12.	Discipline of Children	1.35	0.67	1.44	1.00
13.	Women in today's society	2.68	1.13	2.54	1.24
14.	Family finances	1.65	0.78	1.94	1.19
15.	Careers for women	1.43	0.67	1.55	0.90
16.	Sex relations	2.34	0.95	2.64	1.31
17.	Institution of marriage	3.02	1.54	3.09	1.42
18.	Travelling	1.35	0.64	1.68	0.85
19.	Revising university level education	2.57	1.07	2.69	1.27
20	University	2.08	0.81	1.54	0.68
21.	Third World War	2.64	1.14	2.73	1.12
22.	Inflation	2.08	1.06	2.20	1.15
23.	World's future	1.60	0.86	1.73	0.90
24.	Career-mindedness	2.22	1.08	1.78	0.91

Table 2 continued

	Teacher Ratings	Mean	S.D.
1.	Instructor in general	3.60	0.98
2.	Overall value of course	3.49	1.02
3.	How much learned	3.42	0.89
4.	Discusses points of view	3.45	0.99
5.	Explains clearly	3.71	0.88
6.	Encourages class discussion	3.57	0.88
7.	Genuine interest in students	3.93	0.91
8.	Dynamic and energetic person	3.89	0.90
9.	Interesting style of presentation	3.55	0.94
10.	Contrasts theories	3.30	0.80
11.	Share knowledge and experience with teacher	3.68	0.90
12.	Friendly toward students	4.12	0.87
13.	Well prepared	4.03	0.88
14.	Discusses recent developments	3.58	0.88
15.	Presents origins of ideas and concepts	3.50	0.90
16.	Lectures easy to outline	3.74	0.93
17.	Relates to students	3.89	0.94
18.	Enjoys teaching	4.13	0.82
19.	Gives references for points	3.51	0.96
20.	Enthusiastic about subject	4.14	0.82
21.	Has self-confidence	4.16	0.77
22.	Recognizes and greets students	3.96	0.93
23.	Clarifies thinking	3.53	0.84
24.	Invites criticism	3.32	0.95
25.	Careful and precise answers	3.66	0.85
26.	Summarizes major points	3.69	0.83
27.	Accessible to students	3.81	0.86
28.	Class understands him	3.65	0.98

Table 3

Regression of TRF Scores on Total Similarity With And Without Teacher Effects Removed

Predictor	SS	df	MS	F	p<	ΔR^2
Total similarity	5.99	1	5.99	19.44	.001	.08
Residual	67.78	220	0.31			
			Stepwise Procedure			
Teachers	23.85	13	1.83	7.96	.001	.32
Total similarity	2.17	1	2.17	9.42	.01	.03
Residual	47.75	207	0.23			

Table 4

Regression of TRF Scores on the Set of Individual Similarity Scores and Both Sets of Component Scores Independent of Teacher Effects

Predictor	SS	df	MS	F	p<	ΔR^2
Teachers	23.85	13	1.83	8.32	.001	.32
Similarity	9.62	24	0.40	1.85	.05	.13
Residual	40.29	184	0.22			
			Component Scores			
Teachers	23.85	13	1.83	9.63	.001	.32
Component scores	18.13	48	0.37	1.90	.01	.25
Residual	31.78	160	0.19			

Table 5

Stepwise Regression of TRF Scores on Each Set of Component Scores Independent of Teacher Effects

Predictor	SS	df	MS	F	p<	ΔR^2
		Order A				
Teachers	23.85	13	1.83	9.63	.001	.32
Self-ratings	7.16	24	0.30	1.50	n.s.	-
Perceived ratings	10.97	24	0.46	2.30	.01	.15
Residual	31.78	160	0.19			
		Order B				
Teachers	23.85	13	1.83	9.63	.001	.32
Perceived ratings	11.63	24	0.48	2.56	.001	.16
Self-ratings	6.5	24	0.27	1.36	n.s.	-
Residual	31.78	160	0.19			

Appendix A

Survey of Attitudes

1. Professors and Student Needs

a. I feel that university professors are completely indifferent to student needs.
b. I feel that university professors are indifferent to student needs.
c. I feel that university professors are slightly indifferent to student needs.
d. I feel that university professors are slightly concerned about student needs.
e. I feel that university professors are concerned about student needs.
f. I feel that university professors are very much concerned about student needs.

2. Birth Control

a. I am very much in favor of most birth control techniques.
b. I am in favor of most birth control techniques.
c. I am mildly in favor of most birth control techniques.
d. I am mildly opposed to most birth control techniques.
e. I am opposed to most birth control techniques.
f. I am very much opposed to most birth control techniques.

3. Sports

a. I like sports very much.
b. I like sports.
c. I like sports to a slight degree.
d. I dislike sports to a slight degree.
e. I dislike sports.
f. I dislike sports very much.

4. Money

a. I strongly believe that money is not one of the most important goals in life.
b. I believe that money is not one of the most important goals in life.
c. I believe that perhaps money is not one of the most important goals in life.
d. I believe that perhaps money is one of the most important goals in life.
e. I believe that money is one of the most important goals in life.
f. I strongly believe that money is one of the most important goals in life.

5. Grades

a. I am very much in favor of the university grading system as it now exists.
b. I am in favor of the university grading system as it now exists.
c. I am mildly in favor of the university grading system as it now exists.
d. I am mildly opposed to the university grading system as it now exists.
e. I am opposed to the university grading system as it now exists.
f. I am very much opposed to the university grading system as it now exists.

6. Political Orientation

a. I am strongly liberal in my views.
b. I am liberal in my views.
c. I am slightly liberal in my views.
d. I am slightly conservative in my views.
e. I am conservative in my views.
f. I am strongly conservative in my views.

7. Welfare System

a. I am very much opposed to increased public welfare.
b. I am opposed to increased public welfare.
c. I am mildly opposed to increased public welfare.
d. I am mildly in favor of increased public welfare.
e. I am in favor of increased public welfare.
f. I am very much in favor of increased public welfare.

8. Creative Work

a. I like doing creative work very much.
b. I like doing creative work.
c. I like doing creative work to a slight degree.
d. I dislike doing creative work to a slight degree.
e. I dislike doing creative work.
f. I dislike doing creative work very much.

9. Socialized Medicine

a. I am very much opposed to socialized medicine as it operates in Quebec.
b. I am opposed to socialized medicine as it operates in Quebec.
c. I am mildly opposed to socalized medicine as it operates in Quebec.
d. I am mildly in favor of socialized medicine as it operates in Quebec.
e. I am in favor of socialized medicine as it operates in Quebec.
f. I am very much in favor of socialized medicine as it operates in Quebec.

10. University Education

a. I strongly believe that it is very important for a person to have a university education in order to be successful.
b. I believe that it is very important for a person to have a university education in order to be successful.
c. I believe that perhaps it is very important for a person to have a university education in order to be successful.
d. I believe that perhaps it is not very important for a person to have a university education in order to be successful.
e. I believe that it is not very important for a person to have a university education in order to be successful.
f. I strongly believe that it is not very important for a person to have a university education in order to be successful.

11. Exercise

a. I strongly believe that daily exercise is not important.
b. I believe that daily exercise is not important.
c. I believe that perhaps daily exercise is not important.
d. I believe that perhaps daily exercise is important.
e. I believe that daily exercise is important.
f. I strongly believe that daily exercise is important.

12. Discipline of Children

a. I strongly believe that both the father and mother should participate in disciplining the children.
b. I believe that both the father and mother should participate in disciplining the children.
c. I believe that perhaps both the father and mother should participate in disciplining the children.
d. I believe that perhaps both the father and mother should not participate in disciplining the children.
e. I believe that both the father and mother should not participate in disciplining the children.
f. I strongly believe that both the father and mother should not participate in disciplining the children.

13. Women in Today's Society

a. I strongly believe that women are not taking too aggressive a role in society today.
b. I believe that women are not taking too aggressive a role in society today.
c. I believe that perhaps women are not taking too aggressive a role in society today.
d. I believe that perhaps women are taking too aggressive a role in society today.
e. I believe that women are taking too aggressive a role in society today.
f. I strongly believe that women are taking too aggressive a role in society today.

14. Family Finances

a. I strongly believe that both the man and woman should handle the family finances.
b. I believe that both the man and woman should handle the family finances.
c. I believe that perhaps both the man and woman should handle the family finances.
d. I believe that perhaps both the man and woman should not handle the family finances.
e. I believe that both the man and woman should not handle the family finances.
f. I strongly believe that both the man and woman should not handle the family finances.

15. Careers for Women

a. I am very much in favor of women pursuing careers.
b. I am in favor of women pursuing careers.
c. I am mildly in favor of women pursuing careers.
d. I am mildly opposed to women pursuing careers.
e. I am opposed to women pursuing careers.
f. I am very much opposed to women pursuing careers.

16. Sex Relations

a. I strongly believe that North American society puts too much emphasis on sex relations.
b. I believe that North American society puts too much emphasis on sex relations.
c. I believe that perhaps North American society puts too much emphasis on sex relations.
d. I believe that perhaps North American society does not put too much emphasis on sex relations.
e. I believe that North American society does not put too much emphasis on sex relations.
f. I strongly believe that North American society does not put too much emphasis on sex relations.

17. Institution of Marriage

a. I strongly believe that the institution of marriage needs reform.
b. I believe that the institution of marriage needs reform.
c. I believe that perhaps the institution of marriage needs reform.
d. I believe that perhaps the institution of marriage does not need reform.
e. I believe that the institution of marriage does not need reform.
f. I strongly believe that the institution of marriage does not need reform.

18. Travelling

a. I like travelling very much.
b. I like travelling.
c. I like travelling to a slight degree.
d. I dislike travelling to a slight degree.
e. I dislike travelling.
f. I dislike travelling very much.

19. Revising University Level Education

a. I strongly believe that university level education needs to be revised.
b. I believe that university level education needs to be revised.
c. I believe that perhaps university level education needs to be revised.
d. I believe that perhaps university level education does not need to be revised.
e. I believe that university level education does not need to be revised.
f. I strongly believe that university level education does not need to be revised.

20. University

a. I like university very much.
b. I like university.
c. I like university to a slight degree.
d. I dislike university to a slight degree.
e. I dislike university.
f. I dislike university very much.

21. Third World War

a. I strongly believe that there is a threat of a third world war.
b. I believe that there is a threat of a third world war.
c. I believe that perhaps there is a threat of a third world war.
d. I believe that perhaps there is no threat of a third world war.
e. I believe that there is no threat of a third world war.
f. I strongly believe that there is no threat of a third world war.

22. Inflation

a. I am very much in favor of government intervention to control inflation.
b. I am in favor of government intervention to control inflation.
c. I am mildly in favor of government intervention to control inflation.
d. I am mildly opposed to government intervention to control inflation.
e. I am opposed to government intervention to control inflation.
f. I am very much opposed to government intervention to control inflation.

23. World's Future

a. I strongly believe that there is a need to be concerned about the world's future.
b. I believe that there is a need to be concerned about the world's future.
c. I believe that perhaps there is a need to be concerned about the world's future.
d. I believe that perhaps there is no need to be concerned about the world's future.
e. I believe that there is no need to be concerned about the world's future.
f. I strongly believe that there is no need to be concerned about the world's future.

24. Career-Mindedness

a. I strongly believe that it is important to be career-minded.
b. I believe that it is important to be career-minded.
c. I believe that perhaps it is important to be career-minded.
d. I believe that perhaps it is not important to be career-minded.
e. I believe that it is not important to be career-minded.
f. I strongly believe that it is not important to be career-minded.